

ARSDIGITA UNIVERSITY MONTH 0: MATHEMATICS FOR COMPUTER SCIENCE

This course is concerned with the study of functions. For the first two weeks, we will study calculus, which will involve the study of some fairly complex functions. We will only skim the surface of much of this material. During the second two weeks, we will study linear algebra, which will involve the study of fairly simple functions – linear functions – in much greater detail.

This is the outline of the course notes from September 2000. In the future, the syllabus may be slightly different, and so these notes may be out of date in a year's time.

LECTURE 1: GRAPHING FUNCTIONS. EXPONENTIAL AND LOGARITHMIC FUNCTIONS.

1.1. What is a function? A *function* is a rule or a map from one set A to another set B . A function takes an element of the first set, the *domain*, and assigns to it an element of the second set, the *range*. We will use the following notation:

f, g, h, F, G are all functions,

$f : A \rightarrow B$ is a function which maps $x \in A$ to $f(x) \in B$.

In calculus, we will be interested in real-valued functions $f : \mathbb{R} \rightarrow \mathbb{R}$. We can graph these functions.

1.2. Graphing Functions. We can make a graph of a function $y = f(x)$ by plotting the points $(x, f(x))$ in the xy -plane.

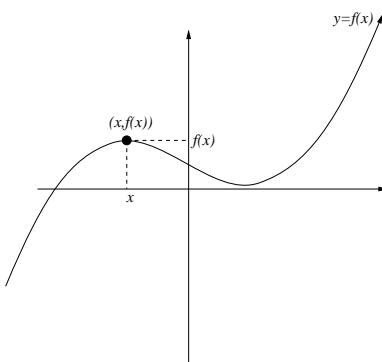


FIGURE 1. Here is how we graph a function $y = f(x)$. We graph the points $(x, f(x))$.

We can make the graph roughly simply by choosing some values, plotting those values, and connecting the dots. We do have to be careful when we choose which x -values to plot.

1.3. Graphing tricks. If we know how to graph a function $y = f(x)$, then we can build up a repertoire of functions by figuring out how this graph changes when we add constants to our equation: $f(x) \rightarrow f(x + c)$, $f(x) \rightarrow f(c \cdot x)$, $f(x) \rightarrow c \cdot f(x)$, or $f(x) \rightarrow f(x) + c$. The graph changes by stretching and shifting, depending on the location of the constant. This allows us to have a whole family of functions whose graphs we can draw, by having a few simple examples of functions in mind.

1.4. Examples of Functions. The first examples are lines. To graph a line, we merely have to plot two points, and connect them with a straight line. The first example we have is the line $y = x$. A general line is of the form $y = mx + b$. The first constant m is called the slope and the second constant b is called the y -intercept.

The next examples are polynomials. These are functions of the form $y = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, where $a_n \neq 0$. We say that the polynomial $y = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ has *degree* n . A line is a polynomial of degree less than or equal to 1.

1.5. Exponentials and logarithms. Let a be some real number greater than 0. Then we can look at the exponential function $y = a^x$ and the logarithmic function $y = \log_a x$.

1.6. Trigonometric functions. There is a family of trigonometric functions, $y = \sin(x)$, $y = \cos(x)$, $y = \tan(x)$, $y = \csc(x)$, $y = \sec(x)$, and $y = \cot(x)$. These are defined for angles x , measured in radians. These functions and their properties will be discussed in recitation.

LECTURE 2: LIMITS AND CONTINUITY. SLOPES AND DERIVATIVES.

This is the beginning of calculus in earnest. Single variable calculus addresses two major problems.

1. Differential calculus addresses the question, what is the slope of the tangent line to $y = f(x)$ at the point $P = (x_0, f(x_0))$?
2. Integral calculus, on the other hand, addresses the question, what is the area under the graph $y = f(x)$ between $x = a$ and $x = b$?

This week, we will look at the first of these questions, and we will apply the answer to many real-world situations.

2.1. Limits. Suppose we have a function $y = f(x)$ defined on an interval around $x = a$, but not necessarily at $x = a$ itself. Suppose there is a real number $L \in \mathbb{R}$ such that as x approaches a (from both the left and the right), $f(x)$ approaches L . In this case, we say,

$$\lim_{x \rightarrow a} f(x) = L \text{ or } f(x) \rightarrow L \text{ as } x \rightarrow a.$$

If no such number exists, then we say there is no limit. We will compute several examples.

2.2. Continuity. We say that a function $y = f(x)$ is continuous at $x = a$ if

1. $f(a)$ is defined;
2. $\lim_{x \rightarrow a}$ exists; and
3. $\lim_{x \rightarrow a} = f(a)$.

We will look at a couple of examples.

2.3. Slope and derivatives. The goal is to compute the slope of the tangent line to a point $(x, f(x))$. We do this by computing the slope of secant lines, and taking the limit as the secant lines approach the tangent lines. Thus, the derivative is

$$\frac{df}{dx} = f'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

We will look at several examples of computing derivatives by this definition.

2.4. Velocity and Rates of Change. In recitation, we will see how derivatives appear in the real world as the velocity of a position function, and as a rate of change. We will look at several problems of this nature.

LECTURE 3: DIFFERENTIATION METHODS.

In this lecture, we will see several methods for computing derivatives, so that we do not always have to use the definition of derivative in order to do computations.

3.1. Rules for differentiation. Here is a list of differentiation rules. We will prove some of these in class, and use them to compute some difficult derivatives.

1. $\frac{d}{dx}c = 0$ for any constant $c \in \mathbb{R}$.
2. $\frac{d}{dx}(x^n) = n \cdot x^{n-1}$.
3. $\frac{d}{dx}(c \cdot f(x)) = c \cdot \frac{d}{dx}(f(x))$, for any constant c .
4. $\frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx}(f(x)) + \frac{d}{dx}(g(x))$.
5. The product rule.

$$\frac{d}{dx}(f(x) \cdot g(x)) = g(x) \cdot \frac{d}{dx}(f(x)) + f(x) \cdot \frac{d}{dx}(g(x)).$$

6. The quotient rule.

$$\frac{d}{dx} \left(\frac{f(x)}{g(x)} \right) = \frac{g(x) \cdot \frac{d}{dx}(f(x)) - f(x) \cdot \frac{d}{dx}(g(x))}{(g(x))^2}.$$

7. The chain rule.

$$\frac{d}{dx}(f(g(x))) = f'(g(x))g'(x) = \frac{df}{du} \cdot \frac{du}{dx},$$

where $u = g(x)$.

8. The power rule.

$$\frac{d}{dx}(u^n) = n \cdot u^{n-1} \cdot \frac{du}{dx}.$$

9. $\frac{d}{dx}(\sin(x)) = \cos(x)$.
10. $\frac{d}{dx}(\cos(x)) = -\sin(x)$.
11. $\frac{d}{dx}(e^x) = e^x$.
12. $\frac{d}{dx}(a^x) = \ln(a) \cdot a^x$.
13. $\frac{d}{dx}(\ln(x)) = \frac{1}{x}$.
14. $\frac{d}{dx}(\log_a(x)) = \frac{1}{\ln(a) \cdot x}$.

We will also discuss implicit functions and implicit differentiation.

3.2. Graphing functions. Derivatives can be used to graph functions. The first derivative tells us whether a function is increasing or decreasing. The second derivative tells us about the concavity of the function. This will be discussed in recitation.

LECTURE 4: MAX-MIN PROBLEMS. TAYLOR SERIES.

Today, we will be using derivatives to solve some real-world problems.

4.1. Max-min problems. The idea with max-min problems is that a function $f(x)$ defined on an interval $[a, b]$ will take a maximum or minimum value in one of three places:

1. where $f'(x) = 0$,
2. at a , or
3. at b .

Thus, if we are given a problem such as the following, we know where to look for maximums, minima, the biggest, the smallest, the most, the least, etc.

EXAMPLE. Suppose we have 100 yards of fencing, and we want to enclose a rectangular garden with maximum area. What is the largest area we can enclose? \diamond

We will discuss methods for solving problems like the following.

4.2. Approximations. We can also use the derivative to compute values of functions. Suppose we want to compute $\sqrt{4.02}$. Then we can take a linear approximation, the tangent line, of \sqrt{x} at $x = 4$, and compute the approximate value of $\sqrt{4.02}$ using the tangent line. Linear and quadratic approximations will be discussed.

4.3. Newton's method. Often times, we would like to know when a function takes the value 0. Sometimes, the answer to this question is obvious, and other times it is more difficult. If we know that $f(x)$ has a zero between $x = 1$ and $x = 2$, then we can use Newton's method to determine the exact value of x for which $f(x) = 0$. This will be discussed in recitation.

LECTURE 5: TAYLOR SERIES. ANTIDERIVATIVES.

In this lecture, we will tie up loose ends of differential calculus, and begin integral calculus.

5.1. Taylor series. The simplest functions that we have seen so far are polynomials. These are easy to evaluate and easy to differentiate. Suppose we have some more difficult function, $y = f(x)$. Then we hope to approximate $y = f(x)$ using a polynomial. Our first guess is the tangent line. We can add on higher terms using higher order derivatives. If we infinitely many terms, we end up with the Taylor series. The computation of the Taylor series will be discussed.

5.2. Antiderivatives. If we are given a function $f(x)$, we say that another function $F(x)$ is an *antiderivative* of $f(x)$ if $F'(x) = f(x)$. We will write

$$\int f(x) \, dx = F(x).$$

Notice that a constant can always be added to $F(x)$, since $\frac{d}{dx}c = 0$! We will compute several anti-derivatives.

Taking antiderivatives can be quite difficult. There is no general product rule or chain rule, for example. We have some new rules, however.

5.3. Change of variables. The change of variables formula is the antidifferentiation version of the chain rule. It can be stated as follows.

$$\int f(g(x))g'(x) \, dx = \int f(u) \, du.$$

There is also an antidifferentiation version of the product rule. We will discuss this at a later time.

LECTURE 6: AREA UNDER A CURVE. FUNDAMENTAL THEOREM OF CALCULUS.

I claimed at the beginning of the course that integral calculus attempted to answer a question involving the area under a curve. We will see how this is related to antidifferentiation today.

6.1. Riemann sums. We can approximate the area between a function and the x -axis using Riemann sums. The idea is that we use rectangles to estimate the area under the curve. We sum up the area of the rectangles, and take the limit as the rectangles get more and more narrow. This leads us to the left-hand and right-hand Riemann sums. The left-hand Riemann sum is

$$\int_a^b f(x) \, dx = \lim_{\Delta x \rightarrow 0} \sum_{k=0}^{n-1} f(x_k) \cdot \Delta x.$$

The right-hand Riemann sum is

$$\int_a^b f(x) \, dx = \lim_{\Delta x \rightarrow 0} \sum_{k=1}^n f(x_k) \cdot \Delta x.$$

These both converge to the area we are looking for, so long as $f(x)$ is a continuous function.

We will make several computations using Riemann sums, and we will look at a few more examples of Riemann sums.

These Riemann sums seem to have to relation to the antiderivatives discussed above, although we have used the same notation. It turns out that these two ideas are related.

6.2. The Fundamental Theorem of Calculus. We will prove the following theorem.

Theorem 6.2.1. *If $f(x)$ is a continuous function, and $F(x)$ is an antiderivative of $f(x)$, then*

$$\int_a^b f(x) \, dx = F(b) - F(a).$$

We will discuss the implications of this theorem, and look at several examples. Furthermore, we will discuss geometric versus algebraic area.

LECTURE 7: SECOND FUNDAMENTAL THEOREM OF CALCULUS. INTEGRATION BY PARTS. THE AREA BETWEEN TWO CURVES.

7.1. Discontinuities and Integration. First, let's talk about one horrific example where everything seems to go wrong. This will be an example of *why* we need a function $f(x)$ to be continuous in order to integrate. Consider the function

$$f(x) = \begin{cases} 0 & x \text{ is a rational number,} \\ 1 & x \text{ is an irrational number.} \end{cases}$$

We can't graph this function. The problem is, between every two rational points, there is an irrational point, and between every two irrational points, there is a rational point. It turns out that this function is also impossible to integrate. We will discuss why.

7.2. The Second Fundamental Theorem of Calculus. When we were taking derivatives, we never really came across functions we couldn't differentiate. From our little bit of experience with taking indefinite integrals (the ones of the form $\int f(x) \, dx$), it doesn't seem obvious that we can always integrate a function. The Second Fundamental Theorem of Calculus tells us that an integral, or antiderivative, always does exist, and we can write one down.

Theorem 7.2.1. *(Second FTC) If $f(x)$ is a continuous function, and a is some real number, then*

$$\frac{d}{dx} \int_a^x f(t) \, dt = f(x).$$

This is a by-product of the proof we gave yesterday of the First Fundamental Theorem of Calculus.

7.3. Integration by Parts. I talked about forming the “inverse” of the Product Rule from differentiation to make a rule about taking integrals. This is known as *integration by parts*, and the rule is

$$\int u \, dv = u \cdot v - \int v \, du.$$

We will discuss this rule further, how to apply it, and some tricks for knowing which choices to make for u and v .

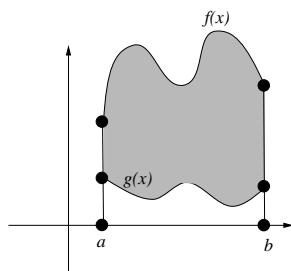


FIGURE 2. We're trying to compute the shaded area between $f(x)$ and $g(x)$ from $x = a$ to $x = b$.

7.4. The area between two curves. Finding the area between two curves looks pretty difficult. Suppose $f(x) > g(x)$. See the figure. If we think about it, though, this is not a hard problem at all! When we think about it, all we have to do is take the area under $f(x)$ on $[a, b]$, and subtract from that the area under $g(x)$ on $[a, b]$. Thus, the area between $f(x)$ and $g(x)$ on $[a, b]$ is

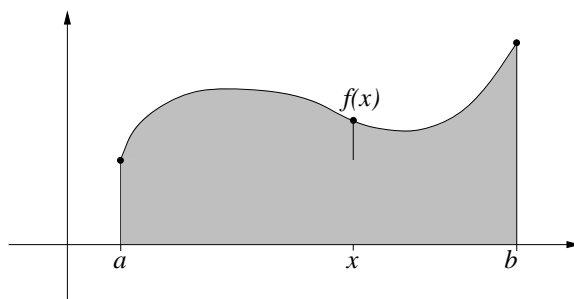
$$\int_a^b f(x) \, dx - \int_a^b g(x) \, dx = \int_a^b (f(x) - g(x)) \, dx.$$

We will discuss some subtleties with this rule.

LECTURE 8: VOLUMES OF SOLIDS OF ROTATION. INTEGRATION TECHNIQUES.

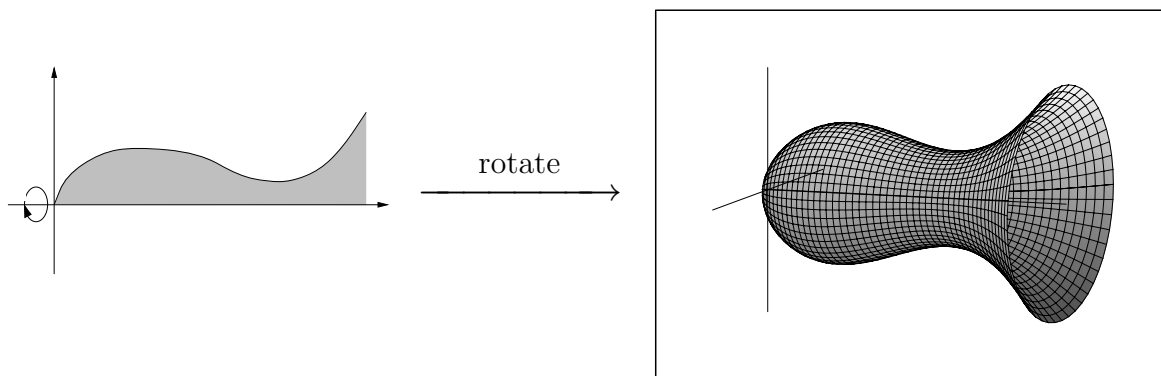
8.1. Volumes of solids of rotation. Today, we'll talk about two methods to compute the volumes of rotational solids, volumes by discs and volumes by cylindrical shells.

Let us consider for a moment the intuitive idea behind a definite integral. The definite



integral $\int_a^b f(x) \, dx$ is supposed to be an infinitesimal sum (denoted by \int_a^b) along the x -axis (denoted by dx) of slices with height $f(x)$. The methods we will present to compute volumes of rotation will have a similar intuitive idea.

8.1.1. Volume by slices. Consider a function $f(x)$ from a to b . We can take the region between this curve and the x -axis and rotate it around the x -axis. We get a *solid of revolution*.



The goal is to compute the volume of this solid. We have seen some examples of this already in recitation. The first method we will use to compute a volume like this is area by

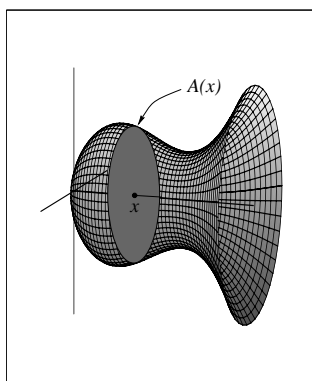


FIGURE 3. This shows the slice at the point x . We say that this slice has area $A(x)$.

slices. The volume is the same as infinitesimally adding up the area $A(x)$ of the slices, as shown in the above figure. So

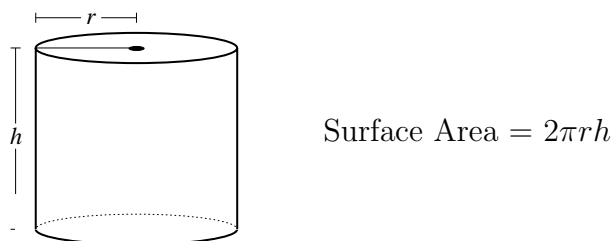
$$\text{Volume} = \int_a^b A(x) dx.$$

But this slice is just a circle with radius $f(x)$, so $A(x) = \pi \cdot r^2 = \pi \cdot f(x)^2$.

$$\text{Volume} = \int_a^b \pi f(x)^2 dx.$$

8.1.2. *Volume by shells.* Again, we consider a solid of revolution. In this case, instead of slicing up our solid, we think of it as filled by cylinders, paper-thin cylinders. Thus, we will often use this method when the solid whose volume we wish to determine is described by a

function rotated around the y -axis rather than the x -axis. We can think of the volume as the infinitesimal sum of the surface area of the cylinders:



Thus, when computing the volume by cylindrical shells, we have

$$\text{Volume} = \int_a^b 2\pi x f(x) dx.$$

Care is required to figure out what the appropriate height function is, and what bounds of integration to use. Furthermore, it is often useful to find a volume such as one of these by integrating with respect to y rather than x . We will see several examples of the cases to which these slices and cylindrical shells apply.

8.2. Integration using partial fractions. A rational function is the quotient of two polynomials, $\frac{P(x)}{Q(x)}$. It is *proper* if

$$\deg(P(x)) < \deg(Q(x))$$

and *improper* if

$$\deg(P(x)) \geq \deg(Q(x)).$$

An improper rational function can always be reduced to a polynomial plus a proper rational function. For example,

$$\begin{aligned} \frac{x^4}{x^4 - 1} &= \frac{(x^4 - 1) + 1}{x^4 - 1} = 1 + \frac{1}{x^4 - 1}, \text{ and} \\ \frac{x^5}{x^4 - 1} &= \frac{(x^5 - x) + x}{x^4 - 1} = x + \frac{x}{x^4 - 1}. \end{aligned}$$

Furthermore, we can “simplify” or combine terms:

$$\frac{2}{x+3} + \frac{6}{x-2} = \frac{2(x-2) + 6(x+3)}{(x-2)(x+3)} = \frac{8x+14}{x^2+x-6}.$$

How can we go backwards? This is most easily illustrated with an example. If we start with

$$\frac{2x-11}{(x+2)(x-3)} = \frac{A}{x+2} + \frac{B}{x-3},$$

how do we find A and B ? We know that A and B must satisfy

$$\begin{aligned} A(x-3) + B(x-2) &= 2x-11, \text{ so} \\ (A+B)x + (2B-3A) &= 2x-11, \end{aligned}$$

so

$$\begin{aligned} A + B &= 2, \\ -3A + 2B &= -11. \end{aligned}$$

We can solve this system, or we can substitute $x = 3, -2$ in $A(x - 3) + B(x + 2)$ to get

$$\begin{aligned} A(3 - 3) + B(3 + 2) &= 2 \cdot 3 - 11 \\ 5B &= -5 \\ B &= -1, \end{aligned}$$

and

$$\begin{aligned} A(-2 - 3) + B(-2 + 2) &= 2(-2) - 11 \\ -5A &= -15 \\ A &= 3. \end{aligned}$$

So

$$\frac{2x - 11}{(x + 2)(x - 3)} = \frac{3}{x + 2} - \frac{1}{x - 3}.$$

What is this good for? Suppose we want to integrate

$$\int \frac{2x - 11}{(x + 2)(x - 3)} dx.$$

We might try to substitute $u = x^2 + x - 6$, but we would soon see that $(2x - 11) dx$ is *not* du . So instead, we can split this up:

$$\begin{aligned} \int \frac{2x - 11}{(x + 2)(x - 3)} dx &= \int \frac{3}{x + 2} dx - \int \frac{1}{x - 3} dx \\ \begin{array}{cc} u &= x + 2 \\ du &= dx \end{array} &\quad \begin{array}{cc} w &= x - 3 \\ dw &= dx \end{array} \\ &= \int \frac{3}{u} du - \int \frac{1}{w} dw \\ &= 3 \ln |x + 2| - \ln |x - 3| + c \\ &= \boxed{\ln \left(\frac{|x + 2|^3}{|x - 3|} \right) + c}. \end{aligned}$$

A couple of cautionary notes about partial fractions:

1. If there is a term $(x - a)^m$ in the denominator, we will get a sum of the form

$$\frac{A_1}{x - a} + \frac{A_2}{(x - a)^2} + \cdots + \frac{A_m}{(x - a)^m},$$

where the A_i are all *constants*.

2. If there is a quadratic factor $ax^2 + bx + c$ in the denominator, it gives a term

$$\frac{Ax + b}{ax^2 + bx + c}.$$

8.3. Differential Equations. Finally, if there is time, we will talk a bit about differential equations. A differential equation is an equation that relates a function and its derivatives. For example, consider

$$\frac{dy}{dx} = xy.$$

The goal is to determine what y is as a function of x . If we treat the term $\frac{dy}{dx}$ as a fraction, we can separate the variables and put all terms with y and dy on the left-hand side, and all terms with x and dx on the right-hand side. Upon separating, we get

$$\frac{dy}{y} = x \, dx.$$

We can now integrate each side to get

$$\ln |y| = \int \frac{dy}{y} = \int x \, dx = \frac{x^2}{2} + c.$$

Thus, by raising each side to the power of e , we get

$$y = C \cdot e^{x^2},$$

where C is a constant. Differential equations that can be solved in this manner are called *differential equations with variables separable*. There are many other differential equations which can not be solved in this way. Many courses in mathematics are devoted to solving different kinds of differential equations. We will not spend any more time on them in this course.

LECTURE 9: TRIGONOMETRIC SUBSTITUTION. MULTIVARIABLE FUNCTIONS.

9.1. Trigonometric substitutions. Today, we'll discuss trigonometric substitutions a little further. To use the method of trigonometric substitution, we will make use of the following identities.

$$\begin{aligned} 1 - \sin^2 \theta &= \cos^2 \theta, \\ 1 + \cot^2 \theta &= \csc^2 \theta, \\ \sec^2 \theta - 1 &= \tan^2 \theta. \end{aligned}$$

If we see an integral with the function $\sqrt{a^2 - x^2}$ in the integrand, with a a constant, we can make the substitution $x = a \sin \theta$. Then we have

$$\sqrt{a^2 - x^2} = \sqrt{a^2 - a^2 \sin^2 \theta} = \sqrt{a^2 \cos^2 \theta} = a \cos \theta.$$

Notice that in the case, $dx = a \cos \theta \, d\theta$. We make similar substitutions in the remaining cases. If we see a function of the form $\sqrt{x^2 - a^2}$ in the integrand, we can make the substitution $x = a \sec \theta$. If we see a function of the form $\sqrt{x^2 + a^2}$ in the integrand, we can make the substitution $x = a \tan \theta$.

9.2. Cavalieri's Principle. Next, we'll discuss Cavalieri's Principle. This relates to the disc method we discussed yesterday, although Cavalieri, a contemporary of Galileo, stated the result centuries before the advent of calculus. We will see that this principle is true using calculus.

Theorem 9.2.1. *If two solids A and B have the same height, and if cross-sections parallel to their bases have the same corresponding areas, then A and B have the same volume.*

This is clearly true, if we take the volumes by discs. Since the areas of the discs are the same, the volume that we compute when we integrate will be the same.

9.3. Multivariable Functions. Finally, there will be a brief introduction to multivariable functions. We will discuss graphing multivariable functions and multiple integrals, as time permits.

Given a multivariable function $z = f(x, y)$, we can graph its *level curves*. These are the curves that we obtain by setting $z = a$, for constants a . As a varies, we get all of the level curves of the function, which we graph in the xy -plane. For example, we can draw the level curves of the function $z = x^2 + y^2$. These will just be circles $a = x^2 + y^2$. Using

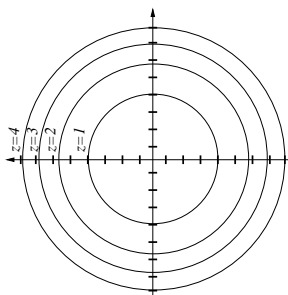


FIGURE 4. This shows the level curves of the function $z = x^2 + y^2$.

the level curves, we can get a three dimensional view of the function. Now instead of a

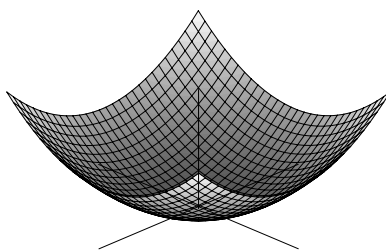


FIGURE 5. This shows the three dimensional graph of the function $z = x^2 + y^2$. This graph is known as a paraboloid.

tangent line to a curve, we have a *tangent plane* to a surface. We can still compute *partial* derivatives. But there are many directions in which we can differentiate. We now have *directional derivatives*. Finally, we can also compute the *volume* under a surface $z = f(x, y)$ over a region R . To do this, we can break R up into many small squares, and add up lots of rectangular solids, taking their height to be some average value of the function. We can

make exactly the same Riemann sums as before, but this time instead of just dx , we have $dA = dx \, dy = dy \, dx$.

LECTURE 10: LINEAR ALGEBRA

Today, we make a total break from calculus and start linear algebra. We (probably) won't take another derivative or integral this month! In linear algebra, we will discuss a special class of functions, linear functions

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m.$$

We will begin by discussing vectors, properties of vectors, and matrices.

10.1. **Vectors.** We have *vectors* in \mathbb{R}^n . For example,

$$\begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \\ 3 \end{bmatrix} \in \mathbb{R}^5.$$

In general, we will think of vectors in \mathbb{R}^n as column vectors, but we may write them as row vectors in order to save space. We will use v and w to represent vectors. We add vectors component-wise. That is,

$$\begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \\ 5 \\ -4 \\ 1 \end{bmatrix} = \begin{bmatrix} 1+2 \\ 2+0 \\ 1+5 \\ 0-4 \\ 3+1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 6 \\ -4 \\ 4 \end{bmatrix}.$$

We can multiply a vector by a real number $a \in \mathbb{R}$, a *scalar*, and we do this component-wise as well. We cannot multiply two vectors, but we can take their *dot product*.

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^n x_i \cdot y_i.$$

The dot product of two vectors is a real number. The length of a vector v is $\|v\| = \sqrt{v \cdot v}$. A *unit vector* is a vector v with $\|v\| = 1$. If we have two vectors v and w in \mathbb{R}^n , then the angle θ between them can be computed by

$$\cos \theta = \frac{v \cdot w}{\|v\| \cdot \|w\|}.$$

This is known as the Law of Cosines. Notice that if the vectors are perpendicular, then the cosine is 0, so the dot product of the vectors must be 0. We also have the Schwartz Inequality, which says

$$v \cdot w \leq \|v\| \cdot \|w\|.$$

Note that this follows from the Law of Cosines.

10.2. Matrices. A matrix is an $m \times n$ array. An $m \times n$ matrix has m rows and n columns. We will use a matrix to store data. The (i, j) entry of a matrix is the entry in the i^{th} row and j^{th} column. Consider the matrix

$$\begin{bmatrix} 342 & 216 & 47 \\ 312 & 287 & 29 \end{bmatrix}.$$

The first row might represent sales in August 2000, and the second sales in August 1999. The first column might represent gallons of milk, the second loaves of bread, and the third heads of lettuce. So long as we remember what each row and column represents, we only need keep track of the entries in each position.

As there were vector operations, we also have matrix operations. We can add two matrices, so long as they are the same size. As before, addition is entry-wise. For example,

$$\begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} + \begin{bmatrix} -2 & 3 \\ 7 & 4 \end{bmatrix} = \begin{bmatrix} 1-2 & 2+3 \\ -2+7 & 0+4 \end{bmatrix}.$$

We can also multiply a matrix by a scalar. Again, this is entry-wise. Finally, we *can* multiply two matrices together, so long as they have the right sizes. We can multiply an $m \times n$ matrix by an $n \times p$ matrix to get a $m \times p$ matrix. The way this happens is if we are multiplying $A \cdot B$ to get C , then the (i, j) entry of C is the dot product of the i^{th} row of A with the j^{th} column of B . Notice that the dimension restrictions on A and B are such that these two vectors are the same size, so their dot product is well-defined. For example,

$$\begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} \cdot \begin{bmatrix} -2 & 3 \\ 7 & 4 \end{bmatrix} = \begin{bmatrix} 1 \cdot (-2) + 2 \cdot 7 & -2 \cdot (-2) + 0 \cdot 7 \\ 1 \cdot 3 + 2 \cdot 4 & -2 \cdot 3 + 0 \cdot 4 \end{bmatrix}.$$

We can also use a matrix to represent the coefficients of linear equations. For example, if we have the system of equations

$$\begin{cases} 342x + 216y + 47z = 1041 \\ 312x + 287y + 29z = 998 \end{cases},$$

we can find values of x , y and z such that both equations are true. If we let $x = 2$, $y = 1$ and $z = 3$, the equations are true. We will use matrices to find all possible solutions of these equations. We can let a matrix hold the coefficients of these equations. We can write down this system of equations using the matrix equation

$$\begin{bmatrix} 342 & 216 & 47 \\ 312 & 287 & 29 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1041 \\ 998 \end{bmatrix}.$$

We will often refer to this form as $A \cdot x = b$, where A is the coefficient matrix, x is the vector of variables, and b is the desired solution vector.

LECTURE 11: GAUSSIAN ELIMINATION. MATRIX OPERATIONS. INVERSES.

Today, we will further explore matrices.

11.1. Gaussian Elimination. First, we will learn an algorithm for solving a system of linear equations. This Gaussian elimination will come up later in the algorithms course, when you have to figure out its time complexity. Suppose you have the following system of equations.

$$\begin{cases} x + 2y - z = 2 \\ x + y + z = 3 \end{cases}.$$

Subtracting the first equation from the second gives a new system with the same solutions.

$$\begin{cases} x + 2y - z = 2 \\ -y + 2z = 1 \end{cases}.$$

This system has infinitely many solutions. We can let the variable z take any value, and then x and y are determined. The existence of infinitely many solutions can also occur in the case of three equations in three unknowns. If a coefficient matrix A is $n \times n$, or *square*, then we say that it is *singular* if $A \cdot x = b$ has infinitely many solutions or no solutions.

Before we give the algorithm for Gaussian elimination, we will define the elementary row operations for a matrix. An *elementary row operation* on an $m \times n$ matrix A is one of the following.

1. Swapping two rows.
2. Multiplying a row by a constant $c \in \mathbb{R}$.
3. Adding c times row i of A to d times row j of A .

Now, given a system of equations such as

$$\begin{cases} x + 2y - z = 2 \\ x + y + z = 3 \end{cases},$$

the *augmented matrix* associated to this system is

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 1 & 1 & 1 & 3 \end{array} \right].$$

In general, for a system $A \cdot x = b$, the augmented matrix is $[A \mid b]$. We say that a matrix is in *row echelon form* when it satisfies the following conditions.

1. All rows consisting entirely of zeroes, if there are any, are the bottom rows of the matrix.
2. If rows i and $i + 1$ are successive rows that do not consist entirely of zeroes, the leading non-zero entry in row $i + 1$ is to the *right* of the leading non-zero entry in row i .

In Gaussian elimination, we perform elementary row operations on the augmented matrix. We perform the operations to both sides of the matrix. We write the vertical bar simply to remind ourselves that the matrix is augmented. The goal of Gaussian elimination is to put the augmented matrix into row echelon form. This then allows us to solve the associated system of equations. We will discuss the precise details in class.

11.2. Inverse matrices. Finally, we will compute the inverse of a matrix. Given a *square* matrix A , its *inverse* is a matrix A^{-1} such that

$$A \cdot A^{-1} = A^{-1} \cdot A = I,$$

where I is the identity matrix. To compute the inverse, we take the augmented matrix $\begin{bmatrix} A & | & I \end{bmatrix}$, and perform elementary row operations on it until it is in the form $\begin{bmatrix} I & | & B \end{bmatrix}$. Then the matrix B is the inverse of A . We will talk about how to prove this in class.

A square matrix is not always invertible. It turns out that a matrix is *not* invertible if and only if it is singular.

LECTURE 12: MATRICES. FACTORIZATION $A = L \cdot U$. DETERMINANTS.

12.1. Facts about matrices. Here are some interesting facts about matrices, their multiplication, addition, and inverses. Each of the following is true when the dimensions of the matrices are such that the appropriate multiplications and additions make sense.

1. It is *not* always true that $AB = BA$!
2. There is an associative law, $A(BC) = (AB)C$.
3. There is also a distributive law $A(B + C) = AB + AC$.
4. $(AB)^{-1} = B^{-1}A^{-1}$.

There are also some important types of matrices. We say that an entry of a matrix is on the *diagonal* if it is the (i, i) entry for some i . A matrix is a *diagonal* matrix if all of its non-diagonal entries are zero. For example, the matrix

$$\begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & 0 \\ 0 & 0 & 0 & \dots & d_n \end{bmatrix}$$

is a diagonal matrix. Diagonal matrices have some nice properties. They commute with each other (that is, $CD = DC$), and they are easy to multiply. The identity matrix I is a special diagonal matrix: all its diagonal entries are 1. Another type of matrix is an *upper triangular matrix*. This is a matrix which has zeroes in every entry below and to the left of the diagonal. A *lower triangular matrix* is a matrix which has zeroes in every entry above and to the right of the diagonal. Finally, the *transpose* of an $m \times n$ matrix A is an $n \times m$ matrix A^T , where the rows and columns are interchanged. In this operation, the (i, j) -entry becomes the (j, i) -entry. If a matrix has $A = A^T$, then we say A is *symmetric*. Notice that $(A^T)^T = A$, and that $(AB)^T = B^T A^T$.

12.2. Factorization of A . We will discuss the factorization

$$A = L \cdot U$$

of a matrix A into a lower triangular matrix L times an upper triangular matrix U . It turns out that we already have the matrix U . This is the matrix that is the result of Gaussian elimination. If we keep track of our row operations, then we can also determine L . Note that if you need to perform a row exchange during elimination, then the factorization will

be of the form $P \cdot A = L \cdot U$. We will discuss how to find L and P if necessary in class. Finally, in the case that A is a symmetric matrix, the factorization $A = L \cdot U$ becomes

$$A = L \cdot D \cdot L^T,$$

where D is a diagonal matrix, and L is lower triangular (and hence L^T is upper triangular).

Our ultimate goal is to factor a matrix A into the form

$$A = C \cdot D \cdot C^{-1},$$

where C is an invertible matrix, and D is a diagonal matrix. This factorization will allow us to compute powers of A easily, since

$$A^n = (C \cdot D \cdot C^{-1})^n = C \cdot D^n \cdot C^{-1}.$$

12.3. Determinants. As the name suggests, the *determinant* of a matrix will determine something about the matrix: if $\det(A) = 0$, then A is a singular matrix. If A is a 2×2 matrix, then

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

and $\det(A) = ad - bc$. We will define the determinant for a larger square matrix recursively. Some important properties of the determinant are the following.

1. $\det(I) = 1$ for I the identity matrix.
2. If two rows of A are the same, the $\det(A) = 0$.
3. If a row of A is all zeroes, then $\det(A) = 0$.
4. If D is a diagonal matrix, then $\det(D)$ is the product of the diagonal entries.
5. If A is upper or lower triangular, then $\det(A)$ is the product of the diagonal entries.
6. $\det(AB) = \det(A) \det(B)$.

LECTURE 13: VECTOR SPACES.

13.1. Application of factorization. We have seen three algorithms thus far. They can be described as follows.

1. Gaussian elimination.
 - (a) $Ax = b$ gives an augmented matrix $\left[\begin{array}{c|c} A & b \end{array} \right]$.
 - (b) Use row operations to transform $\left[\begin{array}{c|c} A & b \end{array} \right]$ into $\left[\begin{array}{c|c} U & c \end{array} \right]$, where U is an upper triangular matrix.
 - (c) Back solve to find solution(s) x .
2. Gauss-Jordan elimination.
 - (a) $Ax = b$ gives an augmented matrix $\left[\begin{array}{c|c} A & b \end{array} \right]$.
 - (b) Use row operations to transform $\left[\begin{array}{c|c} A & b \end{array} \right]$ into $\left[\begin{array}{c|c} I & d \end{array} \right]$.
 - (c) Read off answer(s) $x = d$.
3. $A = L \cdot U$ factorization.
 - (a) $Ax = b$. Just consider A for a minute.
 - (b) Factor $A = L \cdot U$.
 - (c) $Ax = L \cdot U(x) = b$.
 - (i) Forward substitute $Lz = b$.
 - (ii) Back solve $Ux = z$.

We will discuss each of these algorithms, and see why the last algorithm is the fastest.

13.2. Abstract vector spaces. We are not going to take a step towards the more abstract side of linear algebra. We are going to define a *vector space*, something like \mathbb{R}^n . A vector space (over \mathbb{R}) is a set V along with two operations, $+$ addition of vectors and \cdot scalar multiplication, such that for all vectors $v, w, z \in V$ and scalars $c, d \in \mathbb{R}$,

1. $v + w \in V$ and $v + w = w + v$.
2. $(v + w) + z = v + (w + z)$.
3. There is a zero-vector 0 such that $0 + v = v$.
4. For every v , there is an additive inverse $-v$ such that $v + (-v) = 0$.
5. $c \cdot v \in V$ and $c \cdot (d \cdot v) = (c \cdot d) \cdot v$.
6. $c \cdot (v + w) = c \cdot v + c \cdot w$.
7. $(c + d) \cdot v = c \cdot v + d \cdot v$.

Of these, the most important ones are numbers 1 and 5. Most of the other properties are implied by these two. The most important thing to check is that for any two vectors $v, w \in V$, any *linear combination* $c \cdot v + d \cdot w$ of them is also in V .

In class, we will see many examples of vector spaces. We will continue our discussion of vector spaces by defining subspaces. A subset $W \subseteq V$ of a vector space V is a *vector subspace* if it is also a vector space in its own right. A map $L : V_1 \rightarrow V_2$ is a *linear transformation* if $L(c \cdot v + d \cdot w) = c \cdot L(v) + d \cdot L(w)$. The *kernel* or *nullspace* of a linear transformation is the subset $W_1 \subseteq V_1$ such that $L(w) = 0$ for all $w \in W_1$. We write the kernel as $\ker(L)$ or $N(L)$. The nullspace of a vector space is a subspace.

LECTURE 14: SOLVING A SYSTEM OF LINEAR EQUATIONS.

14.1. More vector spaces. First, we will go over several examples of vector spaces. The most important are the *row space* and *column space* of a matrix. Given a matrix A , the row space $R(A)$ is the vector space consisting of all linear combinations of the row vectors of A . The column space $C(A)$ is the vector space consisting of all linear combinations of column vectors of A .

14.2. Solving systems of linear equations. We will discuss the complete solution to a system of linear equations,

$$A \cdot x = b.$$

To find the complete solution, we must put the matrix into *reduced row echelon form*. A matrix is in reduced row echelon form if it satisfies the following conditions.

1. All rows consisting entirely of zeroes, if there are any, are the bottom rows of the matrix.
2. If rows i and $i + 1$ are successive rows that do not consist entirely of zeroes, the leading non-zero entry in row $i + 1$ is to the *right* of the leading non-zero entry in row i .
3. If row i does not consist entirely of zeroes, then its leading non-zero entry is 1.
4. If a column has a leading non-zero entry (a 1!) in it in some row, then every other entry in that column is zero.

The columns that contain a leading entry are called *pivot columns*. The columns that do not contain leading entries are called *free columns*. The rows that are not entirely zeroes

are *pivot rows*. The *rank* of a matrix A is the number of pivot columns (or rows) that it has when it is transformed into reduced row echelon form. We write $\text{rank}(A) = r$. The *pivot variables* are the variables that correspond to the pivot columns. The *free variables* are the variables corresponding to the free columns.

To find a *particular solution* x_p to $Ax = b$, we put $[A \mid b]$ into reduced row echelon form. Then we set the free variables equal to zero, and read off the values of the pivot variables. This is one solution. If a system has free variables, then it has infinitely many solutions. Every solution is of the form $x_p + x_n$, where x_n is a vector in the nullspace of A , that is $Ax_n = 0$. It turns out that we can write down a general solution using the free variables. We will discuss this precisely in class.

14.3. The fundamental subspaces. We will discuss the four subspaces associated to an $m \times n$ matrix A . The four subspaces are the row space $R(A)$, the column space $C(A)$, the nullspace $N(A)$, and the left nullspace $N(A^T)$.

LECTURE 15: LINEAR INDEPENDENCE. BASES.

15.1. Independence and dependence. Spanning sets. Let V be a vector space. Suppose $v_1, \dots, v_n \in V$ is some finite set of distinct non-zero vectors. A *linear combination* of these is a sum

$$c_1v_1 + c_2v_2 + \dots + c_nv_n,$$

where the c_i are real numbers. We say that the set of vectors $\{v_1, \dots, v_n\}$ is *linearly independent* if no non-zero combination of the vectors is zero. That is, if

$$c_1v_1 + c_2v_2 + \dots + c_nv_n = 0$$

for real numbers c_i , then all the c_i must be zero if the v_i are independent.

We say that the set of vectors $\{v_1, \dots, v_n\}$ is *linearly dependent* if there is some non-zero linear combination of them that is zero. This means that one of the vectors is a linear combination of some of the other vectors in the subset. That vector *depends* on the other vectors.

The rank of a matrix can be viewed using the idea of independence. The rank is the number of linearly independent rows (or columns) of a matrix.

We say that a subset of vectors $\{v_1, \dots, v_n\}$ in V *spans* V if every vector $v \in V$ is a linear combination of these vectors. We will also call this subset a *spanning set*.

15.2. Bases and dimension. A basis is a set of vectors $\{v_1, \dots, v_n\}$ in V that is both linearly independent and is a spanning set.

Theorem 15.2.1. *If $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_m\}$ are both bases for a vector space V , then $n = m$.*

We will prove this theorem in class. This allows us to define dimension. The *dimension* of a vector space V is the number of vectors in a basis. It turns out that the dimension of the vector space \mathbb{R}^n is n , so this definition of dimension coincides with the notion we already have. We will prove this in class.

We now return our attention to the fundamental subspaces of a matrix. Suppose that A is an $m \times n$ matrix with rank r . Then it turns out that the dimension of the four fundamental spaces are as follows.

1. The row space $R(A) \subseteq \mathbb{R}^n$ has dimension r .
2. The column space $C(A) \subseteq \mathbb{R}^m$ has dimension r .
3. The nullspace $N(A) \subseteq \mathbb{R}^n$ has dimension $n - r$.
4. The left nullspace $N(A^T) \subseteq \mathbb{R}^m$ has dimension $m - r$.

LECTURE 16: ORTHOGONALITY AND ORTHONORMALITY. GRAM-SCHMIDT ALGORITHM.

16.1. Orthogonality. We say that a pair of vectors v and w in \mathbb{R}^n is *orthogonal* or *perpendicular* if their dot product is zero, $v \cdot w = 0$. In vector spaces that are not \mathbb{R}^n , we must define what we mean by dot product before we can define orthogonal. Thus, we will restrict our attention to $V + \mathbb{R}^n$.

A basis $B = \{v_1, \dots, v_n\}$ of \mathbb{R}^n is an *orthogonal basis* if every pair of vectors is orthogonal. That is, $v_i \cdot v_j = 0$ whenever $i \neq j$. For example, the standard basis of \mathbb{R}^2 is the basis

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

This is an orthogonal basis.

Finally, two subspaces W_1 and W_2 inside V are *orthogonal subspaces* if for every $w_1 \in W_1$ and $w_2 \in W_2$, w_1 and w_2 are orthogonal. Suppose that A is an $m \times n$ matrix with rank r . Then the row space $R(A)$ and the nullspace $N(A)$ are orthogonal subspaces of \mathbb{R}^m , and the column space $C(A)$ and the left nullspace $N(A^T)$ are orthogonal subspaces of \mathbb{R}^n .

16.2. Orthonormality. We say that a basis $B = \{q_1, \dots, q_n\}$ of \mathbb{R}^n is an *orthonormal basis* if every pair of vectors is orthogonal, and if each vector has length 1. That is,

$$q_i \cdot q_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}.$$

The standard basis for \mathbb{R}^2 described above is in fact an orthonormal basis. Orthonormal bases are particularly easy to work with because of the following theorem.

Theorem 16.2.1. *If $B = \{q_1, \dots, q_n\}$ is an orthonormal basis of \mathbb{R}^n and $v \in \mathbb{R}^n$ is any vector, then*

$$v = c_1 q_1 + \dots + c_n q_n,$$

where $c_i = v \cdot q_i$ is the dot product of v and q_i for all i .

We say that a matrix is an *orthogonal matrix* if its columns form an orthonormal basis of the column space of the matrix. We will be particularly interested in square orthogonal matrices. The name orthogonal is a bit of a misnomer. Unfortunately, the name has mathematical roots, and so although we would like to call this an orthonormal matrix, we will not. Suppose

$$Q = \begin{bmatrix} \top & \top & & \top \\ q_1 & q_2 & \cdots & q_n \\ \perp & \perp & & \perp \end{bmatrix}$$

is an orthogonal square matrix. Then Q is invertible, and furthermore, $Q^{-1} = Q^T$.

16.3. Gram-Schmidt algorithm. Suppose $B = \{v_1, \dots, v_n\}$ is any basis of \mathbb{R}^n . We will use the Gram-Schmidt algorithm to turn this into an orthonormal basis $\{q_1, \dots, q_n\}$ such that the subspace spanned by $\{v_1, \dots, v_k\}$ is the same as the subspace spanned by $\{q_1, \dots, q_k\}$ for all k . The algorithm proceeds as follows. First we will turn B into an orthogonal basis. We let

$$w_1 = v_1.$$

Next, the second vector w_2 should be perpendicular to w_1 . We define it to be

$$w_2 = v_2 - \frac{v_2 \cdot w_1}{w_1 \cdot w_1} w_1.$$

We continue in this manner, and let

$$w_i = v_i - \frac{v_i \cdot w_1}{w_1 \cdot w_1} w_1 - \frac{v_i \cdot w_2}{w_2 \cdot w_2} w_2 - \dots - \frac{v_i \cdot w_{i-1}}{w_{i-1} \cdot w_{i-1}} w_{i-1}.$$

At this stage, we could stop, and we would have an orthogonal basis $\{w_1, \dots, w_n\}$ such that the subspace spanned by $\{v_1, \dots, v_k\}$ is the same as the subspace spanned by $\{w_1, \dots, w_k\}$ for all k . To turn this into an *orthonormal* bases, we normalize, and set

$$q_i = \frac{w_i}{\|w_i\|}.$$

This basis $\{q_1, \dots, q_n\}$ is the desired orthonormal basis. We will prove in class that this is indeed an orthonormal basis.

16.4. Change of basis matrices. Let B be the standard basis of \mathbb{R}^n . That is, $B = \{e_1, \dots, e_n\}$, where e_i is the vector that has a 1 in the i^{th} position, and zeroes elsewhere. Let $\tilde{B} = \{v_1, \dots, v_n\}$ and $\hat{B} = \{w_1, \dots, w_n\}$ be two other bases of \mathbb{R}^n . We will now discuss the change of basis matrices. Suppose we have a vector

$$v = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{R}^n.$$

Then we can write this vector as

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1 e_1 + \dots + c_n e_n = I \cdot \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

We may want to write our vector in terms of \tilde{B} . To do this, we need to find constants b_1, \dots, b_n such that

$$v = b_1 v_1 + \dots + b_n v_n = \begin{bmatrix} \top & & \top \\ v_1 & \cdots & v_n \\ \perp & & \perp \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Solving this equation, we get that

$$\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \top & & \top \\ v_1 & \cdots & v_n \\ \perp & & \perp \end{bmatrix}^{-1} \cdot \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Thus, we say that the change of basis from B to \tilde{B} is

$$M_{B \rightarrow \tilde{B}} = \begin{bmatrix} \top & & \top \\ v_1 & \cdots & v_n \\ \perp & & \perp \end{bmatrix}^{-1}.$$

The change of basis matrix from \tilde{B} to B is

$$M_{\tilde{B} \rightarrow B} = \begin{bmatrix} \top & & \top \\ v_1 & \cdots & v_n \\ \perp & & \perp \end{bmatrix}.$$

Finally, if we want to change basis from \tilde{B} to \hat{B} , then the change of basis matrix is

$$M_{\tilde{B} \rightarrow \hat{B}} = M_{B \rightarrow \hat{B}} \cdot M_{\tilde{B} \rightarrow B}.$$

We seem to have multiplied in the wrong order. We will discuss this in class, and compute several examples.

LECTURE 17: EIGENVALUES, EIGENVECTORS AND DIAGONALIZATION.

17.1. Eigenvalues and eigenvectors. For a matrix A , the *eigenvalues* of A are the real numbers λ such that

$$A \cdot v = \lambda \cdot v$$

for some vector v . Let's explore this equation further.

$$\begin{aligned} A \cdot v &= \lambda \cdot v = (\lambda \cdot I) \cdot v \\ Av - \lambda Iv &= 0 \\ (A - \lambda I) \cdot v &= 0 \end{aligned}$$

This last line will have a nonzero solution v if and only if

$$\det(A - \lambda I) = 0.$$

Thus, we find the eigenvalues of A by finding roots of the polynomial

$$\det(A - \lambda I) = 0.$$

An *eigenvector* of a matrix A is a vector v such that

$$A \cdot v = \lambda \cdot v,$$

for some real number λ . Once we know what the eigenvalues are, we can find associated eigenvectors.

Eigenvalues do not necessarily exist, other than 0. They are the roots of a degree n polynomial, and so there are at most n eigenvalues, and there can be no non-zero eigenvalues. If we have n eigenvalues, they may not be distinct. If we have n distinct eigenvalues, then if we choose one eigenvector for each one, these will form a basis for \mathbb{R}^n .

17.2. Diagonalization. If A is a matrix, and there is a diagonal matrix D and an invertible matrix P such that

$$A = P \cdot D \cdot P^{-1},$$

then we say that A is *diagonalizable*.

Theorem 17.2.1. *A matrix A is diagonalizable if it has n distinct eigenvalues. In this case, if $\lambda_1, \dots, \lambda_n$ are the eigenvalues and v_1, \dots, v_n are the eigenvectors, then*

$$A = \begin{bmatrix} \top & & \top \\ v_1 & \cdots & v_n \\ \perp & & \perp \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} \cdot \begin{bmatrix} \top & & \top \\ v_1 & \cdots & v_n \\ \perp & & \perp \end{bmatrix}^{-1}.$$

In the case when the eigenvalues are not distinct, the matrix may or may not be diagonalizable. In this case, there is still a canonical form known as the Jordan canonical form. We will discuss this in class. Some matrices may be diagonalizable over \mathbb{R} , but they are over \mathbb{C} . We will see an application of diagonalization tomorrow.

LECTURE 18: FIBONACCI NUMBERS. MARKOV MATRICES.

18.1. Fibonacci numbers. In the year 1202, the Italian mathematician Fibonacci published a book in which he defined a certain sequence of numbers. The so-called Fibonacci numbers are defined by the relation

$$F_n = F_{n-1} + F_{n-2},$$

and $F_0 = 1$ and $F_1 = 1$. This sequence is

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \dots$$

We can use diagonalization to write a closed form equation for the n^{th} Fibonacci number. If we put two values in a vector,

$$\begin{bmatrix} F_{n-1} \\ F_{n-2} \end{bmatrix},$$

then we can write

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} F_{n-1} \\ F_{n-2} \end{bmatrix} = \begin{bmatrix} F_{n-1} + F_{n-2} \\ F_{n-1} \end{bmatrix} = \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix}.$$

Thus,

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \cdot \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix}$$

So if we could easily compute

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n,$$

we could easily compute the n^{th} Fibonacci number. If we diagonalize the matrix

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

then we can easily compute its n^{th} power. We will do this in class, and come up with the formula

$$F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^{n+1} - \left(\frac{1 - \sqrt{5}}{2} \right)^{n+1} \right].$$

18.2. Markov matrices. We will briefly discuss Markov matrices. This is a place where linear algebra can be applied. A Markov matrix represents a scenario where certain outcome occur with various probabilities, and these probabilities change over time. In some cases, a steady-state is approached, and we will see exactly when. We will use eigenvectors and eigenvalues to determine what that steady-state is.